```
                                                  M  A  A  Q  N  G  N  T        8
AGCCGCAGAGCGCACAGAAAGGAGGCGCCGAGACAGACATCACC ATG GCA GCC CAG AAT GGA AAC ACC     68

 S  F  T  P  N  F  N  P  P  Q  D  H  A  S  S  L  S  F  N  F     28
AGT TTC ACA CCC AAC TTT AAT CCA CCC CAA GAC CAT GCC TCC TCC CTC TCC TTT AAC TTC   128

 S  Y  G  D  Y  D  L  P  M  D  E  D  E  D  M  T  K  T  R  T     48
AGT TAT GGT GAT TAT GAC CTC CCT ATG GAT GAG GAT GAG GAC ATG ACC AAG ACC CGG ACC   188

 F  F  A  A  K  I  V  I  G  I  A  L  A  G  I  M  L  V  C  G     68
TTC TTC GCA GCC AAG ATC GTC ATT GGC ATT GCA CTG GCA GGC ATC ATG CTG GTC TGC GGC   248

 I  G  N  F  V  F  I  A  A  L  T  R  Y  K  K  L  R  N  L  T     88
ATC GGT AAC TTT GTC TTT ATC GCT GCC CTC ACC CGC TAT AAG AAG TTG CGC AAC CTC ACC   308

 N  L  L  I  A  N  L  A  I  S  D  F  L  V  A  I  I  C  C  P    108
AAT CTG CTC ATT GCC AAC CTG GCC ATC TCC GAC TTC CTG GTG GCC ATC ATC TGC TGC CCC   368

 F  E  M  D  Y  Y  V  V  R  Q  L  S  W  E  H  G  H  V  L  C    128
TTC GAG ATG GAC TAC TAC GTG GTA CGG CAG CTC TCC TGG GAG CAT GGC CAC GTG CTC TGT   428

 A  S  V  N  Y  L  R  T  V  S  L  Y  V  S  T  N  A  L  L  A    148
GCC TCC GTC AAC TAC CTG CGC ACC GTC TCC CTC TAC GTC TCC ACC AAT GCC TTG CTG GCC   488

 I  A  I  D  R  Y  L  A  I  V  H  P  L  K  P  R  M  N  Y  Q    168
ATT GCC ATT GAC AGA TAT CTC GCC ATC GTT CAC CCC TTG AAA CCA CGG ATG AAT TAT CAA   548

 T  A  S  F  L  I  A  L  V  W  M  V  S  I  L  I  A  I  P  S    188
ACG GCC TCC TTC CTG ATC GCC TTG GTC TGG ATG GTG TCC ATT CTC ATT GCC ATC CCA TCG   608

 A  Y  F  A  T  E  T  V  L  F  I  V  K  S  Q  E  K  I  F  C    208
GCT TAC TTT GCA ACA GAA ACC GTC CTC TTT ATT GTC AAG AGC CAG GAG AAG ATC TTC TGT   668

 G  Q  I  W  P  V  D  Q  Q  L  Y  Y  K  S  Y  F  L  F  I  F    228
GGC CAG ATC TGG CCT GTG GAT CAG CAG CTC TAC TAC AAG TCC TAC TTC CTC TTC ATC TTT   728

 G  V  E  F  V  G  P  V  V  T  M  T  L  C  Y  A  R  I  S  R    248
GGT GTC GAG TTC GTG GGC CCT GTG GTC ACC ATG ACC CTG TGC TAT GCC AGG ATC TCC CGG   788

 E  L  W  F  K  A  V  P  G  F  Q  T  E  Q  I  R  K  R  L  R    268
GAG CTC TGG TTC AAG GCA GTC CCT GGG TTC CAG ACG GAG CAG ATT CGC AAG CGG CTG CGC   848

 C  R  R  K  T  V  L  V  L  M  C  I  L  T  A  Y  V  L  C  W    288
TGC CGC AGG AAG ACG GTC CTG GTG CTC ATG TGC ATT CTC ACG GCC TAT GTG CTG TGC TGG   908

 A  P  F  Y  G  F  T  I  V  R  D  F  F  P  T  V  F  V  K  E    308
GCA CCC TTC TAC GGT TTC ACC ATC GTT CGT GAC TTC TTC CCC ACT GTG TTC GTG AAG GAA   968

 K  H  Y  L  T  A  F  Y  V  V  E  C  I  A  M  S  N  S  M  I    328
AAG CAC TAC CTC ACT GCC TTC TAC GTG GTC GAG TGC ATC GCC ATG AGC AAC AGC ATG ATC  1028

 N  T  V  C  F  V  T  V  K  N  N  T  M  K  Y  F  K  K  M  M    348
AAC ACC GTG TGC TTC GTG ACG GTC AAG AAC AAC ACC ATG AAG TAC TTC AAG AAG ATG ATG  1088

 L  L  H  W  R  P  S  Q  R  G  S  K  S  S  A  D  L  D  L  R    368
CTG CTG CAC TGG CGT CCC TCC CAG CGG GGG AGC AAG TCC AGT GCT GAC CTT GAC CTC AGA  1148

 T  N  G  V  P  T  T  E  E  V  D  C  I  R  L  K  *              385
```

FIG. 1A

```
ACC AAC GGG GTG CCC ACC ACA GAA GAA GTG GAC TGT ATC AGG CTG AAG TGA                1199

CCCACTGGTGTCACACAATTGAAAACCCCAGTCCAGTACTCAGAGCATCACCCACCATCAACCAAGTTCATAGGCTGCA  1278

TGGGAAATGACATCTGTGTTCATGCCTCCCCCGTGCCCTCAAGAAGCCGAATGCTGCAAAGTCGTAACATACAATGAGA  1357

CTAGACATGAACCAAATCAGCTGACATTTACTGATATCCGCTCGACACCTACTGTGTCCACAATCCCCACAAGGAGATT  1436

AGACACAAGGAGCAGCAACTGACATGGACTGAACATGTACTGTGTGCAAACCACACCAATGAGATTAGACGGGGACAGC  1515

AGGAGCTGACATTTACTCTTCACCTACTGTAATCAAAAACACTTGATTTGATTACAATCAAAAACATATAAAAAACATA  1594

ACAAAGTAGCAGAAGCTATTGGAGTTTCCAAGCTATCTCCAGATATATAGATAGTTCACCCTCCATCTTCCCTAATTCT  1673

GTATCTTACCAGTGCAGGAATATCAAAAGGCTATAGGCCAGGCATGATGGCTCATGCCTGTAATCCCAGCACTTGGGGA  1752

GGCTGAGGCACGTGGATCACTTGAGGTCAGGAGTTCAACCCAGGCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAA  1831

ATACAAAATTAGCTAGGCGTGGTGGCGGGCGCCTGTAATCCCAGTTACTCAGGAGGCTGAAGCAGGAGAATAGCTTGAA  1910

CCTGGGAGTTGGAGTTTGCAGTGAGCTGAGATTGCTCCACTGCACTCCAGCCTGAGTGACAGAGTGAGACTCTGTCTCA  1989

GGAAAAAAACAAACAAACAAACAACAAAACAACAACAACAACAACAACAACCAACGGCTATAGAAGAAGACTCTTAGAC  2068

ACAATGGAAATGTAACGATAAGTTTGTCAGTGCGTGGTTTACAGCATCATGGGAGGTGCGTTACAGCCATCATACTGAA  2147

CTTTCCCACCCACCTCCTACTGCCTCCCAGGGCATTCTCTAGGATTTTGGCTTCAAGAAAAAAAAAAATTCTTATAGTCA  2226

GCCCAGCCTTATGTGGTTATCCACAATGGTGTAATTTCAAAGGAAAGAACCTAAAAATCACTTTCCCACTGATGCTTGA  2305

AAGCTTATCATTTTATTTGGGTGGAGATGGGTAATCCTGAGGTGTCAATTTTTGCCTCCTCAGTGCAAAGGATTTCAGT  2384

GGCTCTGGGGTCAGGGGGAAAGAGGACAGAGAAAAAAGTGGAGGTTGCCACTGGCAATGAACATAATCTCTGTGGGCAT  2463

TTTGCTAAGGACTGGACCACTTTCTAGAACACTCCCTCTTTTACAAAAGGAACTCTACCTAGAATCCAAAGACCTGGGT  2542

TCAGGTCCTAACTCTAAGACTCAAGTCCTAAATTCATGATGTTTTCTCTCTGTGTCTCAGTTTTGCTTTAATGAAATGG  2621

CGATGATGAAAATATCTGCTCTTCATACCTTGCAAGACTGTTGGGAGAGCCCATTGAGGCCATGGTTTGTGAATGTGCT  2700

TTTCAACTGTGCACACGATAAGAATGGAGAAGTGATATTGAACAGTTTATTTGGAGGGAGTTTATTTGGAAACCCCATC  2779

CACTGTGATTTATTAGAGAAATACCCACACTTTTTCATCCCTGTTCTTTGGATGAAAGACTCCTGAAGACTTCACAGTG  2858

TACCTTGTCTACAGTGGGCCAAAAAGGGATCCCTGTTCTTGGTTATAATCTGGGAAATTTAACCTCAGATTCTCAGTGA  2937

CCCCAAGACTCTCAGCATCCCTGCGGTCTTAGAAGTGTTGACAGTCTTCCCTGCATGTTGCAAAATAGCACCCTAGTGC  3016

TGCATAAATATCACTTCTGAATCTGTTTGTATTATTATACATTTGTGGTAACTGTAGGTACACGTCTTCATTTCTTCTT  3095

GATTCATTTTGATGTGGTAGCTATGCAAATGGTACCTGGTTTGGGACTGACCCATCCATATTTGACCAATTCCTAATTT  3174

TTTATAGACAAGGAATTAATTGTTTGCTTGTTTGATTGTTTCTATTATTTGTTGATTTGTTTCTCTGACTGAAGTTTCA  3253

ACCAATGTTTCTTTCTATCACCACCCAGCAGACTCACCTTCAGCCCAATCATTGTACTCTCAGAAAATGCAGGCCGGCA  3332

TGGTGGCTCACATCTGTAATCCCAGCACTTCGGGAGGCCAAGATGGGCAGATCACCTGAGGTCAGGAGTTCAAGACCAG  3411
```

FIG. 1B

CCTGGCCAACATGGCAAAACCCCATCTCTAGAAAAATACAGAAATTAGCTGGCGTGGTGGCACATGCCTGTGGTCCCAG 3490

CTCCTCAGGAGGCTGAGGCATGAGAATTGCTTGAACCCCAGAGGCAGAGGTTGCAGTGAATTGAGATCGCACCACTGCA 3569

CTCCAGCCTGGGTGATAGAGCAAGATTCCATCTCAAAAGGAAAATAAAAGAAAATGCAAACACACTATAATATTAGCCT 3648

AAGCAAAACTGTTAATTCTGATTTACAAAAATTCTTACTTGCTTGGCTTTGAAATGCATTGTGTAATAATGCATTTCAA 3727

AGCCAAGCAAGTAACAATTTTAGGTTATGTACATTTCTATAAATATAATAATTGTATTTTTATTTATTATTCTATCCTG 3806

GCTCTTAGCCGAATCAGGAGATTCTTTAGGAATGGACCATGTACCAGTCAAGTCTGTCAGCAGGATTCATCACCCTGTT 3885

CCTTTTTGTCCTAGAATATACCAACTTCCTTTCATTGAAATTTAACTGAAAAAACTTTTGTAAATATCAGTGTGTATTT 3964

GTGATTTTCCAGTGATTAAAGTGTGATGTTGTTATCCAATTAAATAATTAACATGTGGAATTTAAAAAAAAAAAAAAAG 4043

GGCGGCCGC 4052

FIG. 1C

```
GAATTCCCGGGTCGACCCACGCGTCCGGGCGGCTGGAACTCCCGCTTATTGGTCCCCGGTGGCGATCTTTGGGAGACCA  79

ATAGACGCCCCAGAGGGAGGACACTGGGATCCAGACTCCACTGGAACCCCGCTTTTCAGATCCTGGATGGTATCTGTTC  158

TCCCTAAGGATTCTAACAGGGACCTGCACTCACTGACCCCAGCAGAAGTGCTGAACTCCACGTGAGCGCATCTCCCTGA  237

                                    M   G   P   Q   N   R   N   T   S   F   A   P   12
TACACACCAGCCCACCTGTAGCATCATCAAC ATG GGA CCC CAG AAC AGA AAC ACT AGC TTT GCA CCA  304

 D   L   N   P   P   Q   D   H   V   S   L   N   Y   S   Y   G   D   Y   D   L   32
GAC TTG AAT CCA CCC CAA GAC CAT GTC TCC TTA AAC TAC AGT TAT GGT GAT TAT GAC CTC  364

 P   L   G   E   D   E   D   V   T   K   T   Q   T   F   F   A   A   K   I   V   52
CCC CTG GGT GAG GAT GAG GAT GTG ACC AAG ACA CAG ACC TTC TTT GCA GCC AAA ATT GTC  424

 I   G   V   A   L   A   G   I   M   L   V   C   G   I   G   N   F   V   F   I   72
ATT GGC GTG GCA CTG GCA GGC ATC ATG CTG GTC TGC GGC ATT GGC AAC TTT GTC TTC ATT  484

 A   A   L   A   R   Y   K   K   L   R   N   L   T   N   L   L   I   A   N   L   92
GCT GCC CTC GCC CGC TAC AAG AAG CTG CGC AAC CTT ACC AAC CTC CTC ATT GCT AAC CTG  544

 A   I   S   D   F   L   V   A   I   V   C   C   P   F   E   M   D   Y   Y   V   112
GCC ATC TCT GAC TTC CTG GTG GCG ATC GTC TGC TGC CCC TTT GAG ATG GAC TAT TAT GTA  604

 V   R   Q   L   S   W   A   H   G   H   V   L   C   A   S   V   N   Y   L   R   132
GTA CGG CAG CTT TCC TGG GCG CAT GGT CAC GTG CTT TGT GCC TCC GTC AAC TAC CTT CGT  664

 T   V   S   L   Y   V   S   T   N   A   L   L   A   I   A   I   D   R   Y   L   152
ACG GTC TCC CTG TAC GTC TCC ACC AAC GCT CTG CTG GCC ATC GCT ATT GAC AGA TAC CTC  724

 A   I   V   H   P   L   K   P   R   M   N   Y   Q   T   A   S   F   L   I   A   172
GCT ATT GTC CAC CCT TTG AAA CCA CGG ATG AAT TAT CAG ACC GCT TCC TTC CTG ATC GCT  784

 L   V   W   M   V   S   I   L   I   A   V   P   S   A   Y   F   T   T   E   T   192
TTG GTC TGG ATG GTC TCC ATC CTC ATC GCT GTC CCA TCT GCC TAC TTC ACC ACA GAA ACC  844

 I   L   V   I   V   K   N   Q   E   K   I   F   C   G   Q   I   W   S   V   D   212
```

FIG. 2A

ATC CTC GTT ATC GTC AAG AAT CAA GAA AAA ATC TTC TGT GGT CAG ATC TGG TCG GTG GAC 904

Q Q L Y Y K S Y F L F V F G L E F V G P 232

CAG CAG CTC TAC TAC AAA TCC TAC TTC CTC TTC GTC TTC GGG CTT GAG TTC GTG GGT CCC 964

V V T M T L C Y A R I S Q E L W F K A V 252

GTG GTC ACT ATG ACC CTG TGC TAT GCC AGG ATC TCC CAA GAG CTC TGG TTC AAG GCT GTA 1024

P G F Q T E Q I R K R L R C R R K T V L 272

CCT GGC TTC CAG ACG GAG CAA ATC CGC AAG CGG CTG CGT TGC CGC CGC AAG ACA GTG CTA 1084

L L M G I L T A Y V L C W A P F Y G F T 292

CTG CTC ATG GGC ATC CTC ACA GCC TAC GTG CTG TGC TGG GCG CCG TTC TAT GGC TTT ACC 1144

I V R D F F P T V V V K E K H Y L T A F 312

ATA GTG CGA GAC TTC TTC CCC ACG GTA GTT GTG AAG GAG AAG CAC TAC CTC ACC GCC TTC 1204

Y V V E C I A M S N S M I N T I C F V T 332

TAC GTC GTG GAG TGC ATT GCC ATG AGC AAC AGC ATG ATC AAT ACT ATA TGC TTC GTG ACG 1264

V K N N T M K Y F K K M L R L H W R P S 352

GTC AAG AAC AAC ACC ATG AAA TAC TTC AAG AAG ATG CTG CGG CTC CAC TGG CGG CCC TCT 1324

H Y G S K S S A D L D L K T S G V P A T 372

CAC TAC GGG AGT AAG TCC AGC GCT GAC CTC GAC CTC AAA ACC AGC GGG GTG CCT GCC ACT 1384

E E V D C I R L K * 381

GAA GAG GTG GAT TGT ATC AGA CTA AAG TAG 1414

CCTTCAGGTGTTGCCCAAGGAAAAATTTAACATTCGGTACTCAGTAAATCACACACCATCAACCACTCACAAGCTACAT 1493

GGAAAGATACGGCTGTATTCACGTTCTCCTGCTCTAATGTATCAGGACGCTTCTATGTAATAACATACAGCACAACTGA 1572

TGTCTGCATAACATCTTAGAAGGCAGACACAAATAGTAACAAGTGATGTGGACTGAATGCTTCTGTCTGCAAACCACAC 1651

CAACCAATTATTCAAGGACAAGAGCTGACATGTGAGAATTACCTGCTATGTGCAAAAACAAGTTACCCCCCCAAAAAAT 1730

GATAGAAGCTATTTGGAGTTATTCAGCTCTATCTATCTATCTATCTATCCATCCATCCATCCATCCATCCAGGTCACTA 1809

GAAAGAAGTCACAAATGACTAGCCAGAGTCATGCTACATATTCTTTCATTCTGTATCTTTTCTGCACAGAACTGTCAAA 1888

FIG. 2B

GGCAATAGAATAAAGCACCTAGACATACTAGAAATGTAAGGATAACTCCATCAATAGGGAGACCAAGGCCTCATAGGAA 1967

GAGGGTCCATATAGTATACTGACTTTCCCCACTCCACACCAGTTATCTCCTTAGATATTCTGTACTTATCTGCAATGTT 2046

GTAATTTCAAATGAGGAAAAATAAGGGGACAGGCTTTACCACAGATGTATCAAATCTCATCAAGCCCATAGGGCAAAGA 2125

TGGGAGGCTCCTGACACAAGAAATGTATCCAGTTCTGGATAACTTTAATGCCAAGCATTTCAGGGCTCTGGGGTCTTGG 2204

AGGAAGAGGACACAGAAAGAGCCGAGGTTTCCAGTGGCAATGAGTATAATCTGTCCATTTGCTATGATTTGGACAATTT 2283

TCTAGAACATACTCCGACTTACAAAAGGAACTCTACTTGAGATCCAAAGATCCGGGTAAAAGTCCTAACCCCAGGACTC 2362

ATCTCTGTGTGTCTCCACTGTAATGAAATGGAAATAATGAAAACGGATCATTAGGAACATCAGCCCGGCGAAGTCATGG 2441

TGTGGATGTGATTTTCACCTCTTCCTTTGTGAAGAATGAGGTCGTGAAAAGCTCATTAGAGGGAGTTTGGAATGGAGAA 2520

ACAGCTCCACACTTTTCATCCCTCTTCTTTGAATCGGAGACCACTAAACGCATCTTTGAAGTAGCGTATCTATAGTGAG 2599

GCATAAAGGTCTCCCTGTCACAGAGTGCAATCAAGAAAATACAGTCAATGCCCATACCCTCAGCATCCCTGTGGTCTTA 2678

GACAGTCTTCCCAACAAAGCACTGGTGGACCCCAGGACTGAATTCACTTGTATTATTATGTCATCTACTGAATACTAGG 2757

GTTGATCAAGTTGGCTAGATAGGTATTTTCTTCCTCCTTCACAACCCCATATGTATCCCTCCCTTAAATCCAGTTACTA 2836

AGGAAGACCTTCTTAAACACAGGAGAACCATTATTCTGTCCAGGACACAAATAACCTCTCCAGTAGACACTGTACCCTT 2915

CACATGTCAACAGAATTTGCCTCCTTCTTGTATTTAAACATATCATCCTCCTTTCATTTAGATTTAACCAGAAACCATT 2994

CCTGTAAATTTCAATGTGTTTGTGATACCGCACTGTAAAAAGCGTATGCTGTTATCATATGGAATAATTAACATACAGA 3073

ATTGTAATCGTAGTTCCCAAAAGGTTCCCTACTCCTGTTGTATCTTATGTTTATATGTTTGATGTAAATGGAGCTGTGT 3152

AGCTGTCTAAGCAGCTCAAGCCTGAAATGAGGGAATGTCCAATGGTGTTCTTAGAGCAGGGCCATCTCAGGCTAGCAGC 3231

TGGCCTCAGTCTGTGCTCTCTCGGGAGTGTGTTCTTAAATATGAATTAGCAGCAAACCATTAAAAAAAAAAAAAAAGGG 3310

CGGCCGC 3317

FIG. 2C